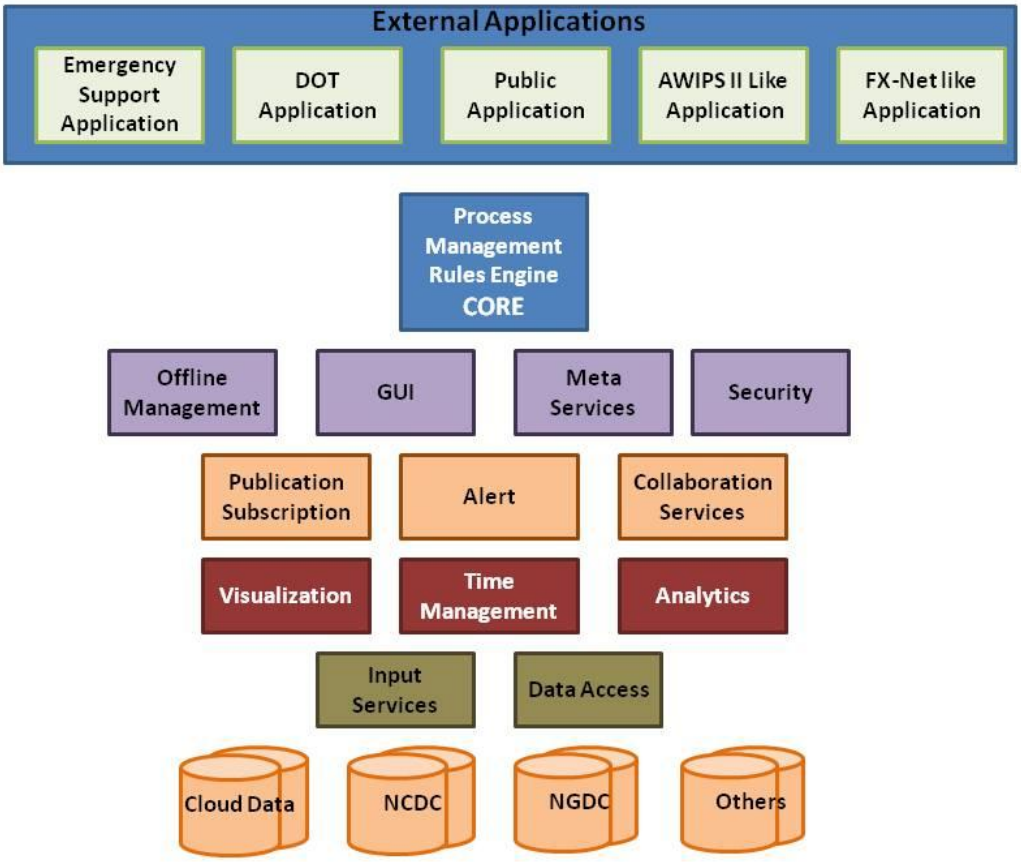
NEIS Concept in Depth



NEIS System

NEIS Concept

In our concept of NEIS, all services are designed to be flexible and adaptable. A service can either pull data into a cloud for processing, or the functions and processing can be pushed out to the edge closer to the data. This allows an infrastructure that can take advantage of a variety of data sources and attempt to minimize the amount of information transferred.

Services can be accessed either through the Process Management Work Flow Engine or directly. By allowing direct access, specialized applications can be built to meet certain performance or specification requirements

To make this model work, the NEIS development team needs agree on a common data exchange model. This data model will be used to exchange data between services. Service will be at least required to support this model, but can server other data formats if necessary.

The NEIS team also recommends using REST interface for all services. More information on REST is available here.

Process Management Workflow Engine (PMWE)

The Process Management/Workflow Engine is envisioned as the main entry point to NEIS System. It guides and manages all aspects of a request allowing a central point to access all NEIS services. All requests are sent through and handled by this module. Services can be directly sent to services if an optimized or specialized application requires it.

The PMWE will have the capability to cache requests. If identical requests are sent through the PMWE, it can verify and respond with the cached version. This takes load off services on the edge.

This is the brains of an NEIS system. It understands how the different components work together and coordinates the work flow of a request. This part can be elastic and is well suited for a cloud environment. Because of this knowledge, requests are redirected to the proper services and location to optimized performance for example:

- A Metadata service request → is redirected to → Metadata Services
- Data Access Request → is redirected to → Data Access Services
- Data Processing Request →
    - data is all from single source → is redirected to → Data Processing Services co-located with data.
    - data is from several sources → is redirected to → Central Data Processing Services which collects data from sources.

By using a single point of entry, the NEIS system can collect a variety of statistics on data usage and patterns. For example:

- What search keywords were used to find the data (i.e. most popular keywords for your data)
- What time of day, day of week is your data being used.
- GUI dependent information (i.e. an emergency manager always uses these types of data)
- What popular keywords don't return results (i.e. areas in need of attention)
- Data access failures (i.e. reliability of services)

This information can then be used in a variety of ways:

- Help improve visibility, usage, performance of data.
- Notify providers of data outages or problems accessing data
- Auto adapting GUI (pushing commonly used data to top of access lists)
- Data suggestions based on similar data and requests. Similar to how Amazon and Netflix make suggestions.

Metadata Services

Existing Metadata services contain a useful library of information. Some are specialized to a particular industry or environment. Others contain information for a variety sources. However,

the NEIS team has noticed each system doesn't quite have all the capabilities to make NEIS concept successful. A couple of key problems include:

- Out of date information - Many existing metadata services contain old/stale non relevant information
- Machine Interface - A way to dynamically discover new relevant data, additional information of where to gather additional information without manual intervention.
- Too much information - A search for popular term like wind could return a couple thousand results, how does a user know which one to use?
- Spatial and Temporal Queries - Some existing Metadata services are moving this way but it is not universal.
- Dynamic information - Based on what data is available, it could be possible to dynamically generate new data. How do you control this type of information?

Metadata services of NEIS would expand upon existing Metadata services and address these issues. The NEIS metadata services would:

- Harvest or federate data from other existing metadata services.
- Provide a metadata dashboard system (discussed below)
- Provide metadata which can be queried both spatially and temporally.
- Provide machine interface to assist dynamic discovery
- Analyze metadata to ensure services and/or links to additional information are alive and active
- Ontology mapping. Different groups use different ontology to refer to the same data.
    - Example (Public consumer versus an Emergency Responder)
- Dynamic metadata population
- Understand potential derivations of data
- Knowledge of what services (valid operations) can be performed on data.
- Automated harvesting
    - As new data comes online, service can automatically harvest relevant metadata and determine what analytics (derived variables and or functions can be performed
- Metadata lineage - what changed, by who, and when.

Additionally, Metadata services would work with the PMWE and allow automatic population of other gathered information such as:

- Data usage and availability statistics
- Keyword relevance (i.e. what keywords are people using to find particular data)
- Common data usage patterns (what other data is being used simultaneously with this data)

Metadata Dashboard

A Metadata Dashboard creates a feedback system to data providers. The idea is provide and easy interface to allowing data providers to quickly see what services their data is compatible

with. Also by providing comparison and feedback, motivate users to provide more metadata information for system.

- Dashboard system would provide the following information and services:
- Automated services underneath to verify data exists, uptime and other statistics.
- Allows providers to see what services their data can interact with based on the metadata provided. (i.e. service compatibility)
- Suggest metadata enhancements to allow data to be used by other services.
- Can rank/score metadata, show comparison against similar sets.
    - This will provide incentive to providers to add/update existing metadata.
- Analysis of services (Uptime over last 4 weeks 99%, last 24 hours 75%)
- Alerting capable for changes on metadata. (Change of units, Broken, change of service url)
- Capability to relay information or news to data users, such as scheduled downtime.

Offline Management

If all NEIS services are web enabled, clients will need a way to take services and offline so similar if not exact functionality is available in an area without Internet connectivity.

Offline Management services would assist clients to take data and services offline (i.e. local to client). It is important to note that it is not just offline data but also services and capabilities.

Analytics

Analytics service would provide processing capabilities for the NEIS system. Many analytics type requests would combine different data to create one response. Analytics services would be created agile and flexible. The means:

- An analytic service request can pull data into NEIS 'cloud' to perform processing.
- An analytic service request can be pushed to data source, eliminating unnecessary data transfer.
- An analytic service request could gather data closest to largest data source, eliminating large data transfers.
- An analytic service would contain many predefined equations.
- An analytic service would work with Metadata services to provide list of possible operations.
- An analytic service would work with PMWE and Metadata services to maintain list of valid operations.
- A user can define a new equation on the fly.

Defined logic and algorithmic equations as a service allowing a user to process and analyze data. Equations are based on ontology and take data (point, array) and provide results.
Example services:

- QC
- Verification
- Trajectories
- Circulation Diagnostics

- Decimation/Downscaling
- Derived products
- Basic Logic Operations (+-*/ Union/Intersect, Max, Min)
- Coordinate re-projection

Specific Examples
- Vector_Velocity = SQRT(Vector_U^2 + Vector_V^2)
- Ocean Example
    - Fish Mortality Functions Salinity of Water vs. Fish Mortality

Time Management

The Time Management service understands how to manage time synchronization allowing user to request data for same times as another existing data set, similar to how D2D Handles time today. This service would be the basis for creating temporally synchronized products

Time Management would need to understand how to sync:
- Observations at specific times
- Forecasts valid at specific times
- Historical products
- Seasonal/Monthly/Yearly (expansive time period) products

Visualization

One concept is Visualization as a Service. This service would allow a consumer to change representation of data. Similar to analytic services, Visualization services would be created agile and flexible. The means:

- A Visualization service request can pull data into NEIS 'cloud' to perform processing.
- A Visualization service request can be pushed to data source, eliminating unnecessary data transfer.
- A Visualization service request could gather data closest to largest data source, eliminating large data transfers.
- A Visualization service would contain many predefined visualizations.
- A Visualization service would work with Metadata services to provide list of possible operations
- A Visualization service would work with PMWE and Metadata services to maintain list of valid operations.
- A user can define a new visualization technique on the fly.

Information about preferred way of visualizing data is something we should store in metadata.

Visualization service would provide the capability to dynamically create:
- Contours
- Vectors
- Polygons
- Rasters
- Points
- Stylized Points (METAR)

Input Services

Some data providers do not have the resources to provide their data in a reliable/robust method. Additionally data providers may not be familiar with data services

Input Services would allow data providers without many resources to make their data available to NEIS. Either through entry forms, or capability to upload compatible data formats.

Additionally Input Services would have a mechanism allowing providers to update metadata information regarding services and/or data.

GUI

A concept of providing GUI functionality as a service. The NEIS GUI module would provide and agile and extensible set or collection libraries and display elements as well as a main web based GUI that combines all elements of NEIS providing a showcasing NEIS capabilities. This module approach would allow other consumers to easily assemble own GUI's.

GUI Services and capabilities would be extensible allowing a developer to specifically implement a requested interface.

Suggested Exploration Areas:

- HTML5 and Video capabilities are very powerful, is it possible to create a streaming video based client.
- The next generation of workers will have more experience with 3D visualization. Because they already have this familiarity, an investigation into new 3D visualization interfaces is probably worthwhile.

Security

Security should not be a standalone service but should be present in every service and is listed as its own module to capture the importance.

Security needs:

- Handles restricted services and/or data.
- Needs to work with Process Management Workflow Engine to ensure security.
- Possibility of using certificates of some kind.

Alert/Messaging Services

Messaging services provide a framework for relaying various messages through the NEIS system. Alerts are one type of message.

Alert/Messaging mechanism should provide capability to allow filtering alerts based on provider, key words, or other metadata.

Alert/Messaging Examples/Categories:

- Alert based on selection criteria
    - Winds over certain speed
    - temperature below value
- Alert based on combinations

    - Winds over certain speed, Humidity below level
    - Ecology data set combined with wind or temperature conditions
- Alert based on new information available in area
    - New observations from boat in certain area
    - Updated forecast over area
- Documentation/Video/Multimedia Alert
    - New research document, video, picture, audio clip in geospatial area
- Metadata alerts
    - Change in service URL's
    - Change in metadata information
    - outage alerts
- Service problems (QoS)
- New data violates or impacts criteria

Publish/Subscription

A Publish/Subscription service would provide capabilities allowing users to subscribe to updates to certain data. Subscription service should provide capability to allow filtering on metadata, spatial, or temporal information.

Collaboration Services

Provide capability for users to collaborate. Collaboration will most likely use existing messaging system described above.

Collaboration will enable the ability to share:

- Voice/Talk
- Text
- Graphics
- Data
- Canned Presentations
- Live Presentations
- Shared Sessions
- Results of annotations tools and utilities (From GUI servers)

Data access

Methods of Data Accessing:

1. Data already exists in the cloud:
    - Either services exist collocated to data ex (NESCC)
    - Or data was uploaded to cloud.

2. A data center implements compatible basic data access services to allow data to be pulled into services of NEIS Cloud (web folder, OGC services, To be defined)

3. A data center implements suite of NEIS Services (Services listed below)

4. Services pushed to edge at data center, i.e. NEIS suite of services are running on hardware collocated at data center.

Document/Multimedia Management

Vast research resources out there in NOAA central library and other locations
Bring spatial search for Research documents, Photos, Multimedia to system.
Require spatial and temporal metadata information for all products.
Provide a link to alert system so users can subscribe to new media in spatial area.